

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12Q 1/68</b>		A1	(11) International Publication Number: <b>WO 97/42345</b> (43) International Publication Date: 13 November 1997 (13.11.97)
(21) International Application Number: PCT/GB97/01163 (22) International Filing Date: 29 April 1997 (29.04.97)  (30) Priority Data: 9609441.2 4 May 1996 (04.05.96) GB		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(71) Applicant ( <i>for all designated States except US</i> ): ZENECA LIMITED [GB/GB]; 15 Stanhope Gate, London W1Y 6LN (GB).  (72) Inventors; and (75) Inventors/Applicants ( <i>for US only</i> ): WHITCOMBE, David, Mark [GB/GB]; Alderley Park, Macclesfield, Cheshire SK10 4TG (GB). LITTLE, Stephen [GB/GB]; Alderley Park, Macclesfield, Cheshire SK10 4TG (GB). BROWNIE, Jannine [GB/GB]; Alderley Park, Macclesfield, Cheshire SK10 4TG (GB).  (74) Agent: PHILLIPS, Neil, Godfrey, Alasdair; Zeneca Pharmaceuticals, Intellectual Property Group, Intellectual Property Dept., Mereside, Alderley Park, Macclesfield, Cheshire SK10 4TG (GB).		Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(54) Title: METHOD FOR DETECTING A NUCLEIC ACID BASE SEQUENCE			
(57) Abstract <p>A method for the detection of diagnostic base sequences in sample nucleic acid. Diagnostic primers having a non-complementary tail comprising tag and detector regions for use in the method. The method is of particular use in combination with the Amplification Refractory Mutation System (ARMS) for the detection of variant diagnostic base sequences against a background of normal diagnostic base sequences.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	RS	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

**METHOD FOR DETECTING A NUCLEIC ACID BASE SEQUENCE.**

The invention relates to novel methods for the detection of diagnostic base sequences in sample nucleic acid. In particular the invention relates to the use of tailed 5 primers in such methods.

The invention is an improvement on currently established procedures for the detection of nucleic acid sequences. The detection of nucleic acid sequences is a desirable and necessary procedure in the following exemplary areas; detection and diagnosis of alleles responsible for genetic diseases in humans and other species; detection and diagnosis of DNA 10 sequences associated or linked to genes that may or may not be involved in disease in humans and other species; detection and diagnosis of neoplasms and the effects of therapy of neoplasms; detection of and distinction between different pathogens (eg. viruses, bacteria and fungi); determining the purity of animal strains and pedigrees; distinguishing and identifying different humans and animal samples in forensic medicine.

15       The polymerase chain reaction (PCR) as disclosed for example in U.S. Patents Nos 4,683,202 and 4,683,195 has been used to amplify specific DNA sequences. However, PCR does not, by itself, provide a method to detect single base mutations. It has been necessary to combine the PCR with other techniques, for example allele specific oligonucleotide probing of PCR amplification products.

20       We have now devised a novel assay system for the detection of diagnostic base sequences which uses tailed diagnostic primers having a tag region and a detector region. Under appropriate conditions any diagnostic primer extension product acts as a template for extension of a further primer. In which case a sequence complementary to the tag region and the detector region will arise in the further primer extension product. A tag primer is provided 25 which can hybridise to the complement of the tag region in the further primer extension product and be extended. A diagnostic base sequence is identified by reference to the sequence complementary to the detector region in the tag primer extension product.

Therefore in a first aspect of the present invention we provide a method for the detection of a diagnostic base sequence in nucleic acid comprised in a sample, which method 30 comprises contacting the sample under hybridising conditions and in the presence of appropriate nucleoside triphosphates and an agent for polymerisation thereof, with a

diagnostic primer for the diagnostic base sequence, the diagnostic primer having a tail sequence comprising a tag region and a detector region, such that an extension product of the diagnostic primer is synthesised when the corresponding diagnostic base sequence is present in the sample, no extension product being synthesised when the corresponding diagnostic base sequence is not present in the sample and any extension product of the diagnostic primer acts as template for extension of a further primer which hybridises to a locus at a distance from the diagnostic base sequence, and contacting the sample with a tag primer which selectively hybridises to the complement of the tag sequence in an extension product of the further primer and is extended, and detecting the presence or absence of the diagnostic base sequence by reference to the detector region in the further primer extension product.

The detector region in the further primer extension product may be detected in a number of ways. For example the sample may be contacted with detector species capable of emitting a detectable signal upon interaction with the detector region in the further primer extension product whereby the presence or absence of the diagnostic base sequence is detected by reference to the detectable signal. It will be appreciated that the detector species cannot become associated with the corresponding detector region until target dependent hybridisation and further primer extension has occurred. This system is well suited for homogeneous assays and real time or end point analysis. A detector species is any species capable of selective association with the detector region in a further primer extension product and release of a detectable signal. It will be appreciated that by "selective association" we mean that the detector species identifies and binds to the detector region in the further primer extension product to the exclusion of other nucleic acid sequences in the sample. Such detector species may include antibodies and hybridisation probe(s). A particular detector species is a detector probe such as a labelled hybridisation probe. Label is conveniently released by the action of for example an exonuclease associated with the polymerase mediated extension of a tag primer. In the specific description hereinafter we describe a number of alternative systems. These include detection of the change in shape of a probe upon hybridisation, the use of two or more probes having interactive labels such as for example the use of fluorescence resonance energy transfer, the use of scintillation proximity assays (SPA), the measurement of a change in fluorescence polarisation upon hybridisation of a fluorescently labelled probe. Further systems will be apparent to the scientist of ordinary skill. These

include the use of a solid phase capture probe for the detector region in the further primer extension product. It will be appreciated that both direct and indirect labelling methods may be used to detect the immobilised further primer extension product. By way of example a further labelled probe for a region other than the detector region may be used. Alternatively, 5 intercalation may be used to detect the detector region/probe DNA duplex. Also, labelled dNTPs may be incorporated into the further primer extension product.

The sequence of the detector probe need not be the same but is conveniently identical to the sequence of the detector region in the tail. It will be appreciated that minor changes may be made to the sequence of the detector probe without affecting its performance 10 to any significant extent.

Alternatively the complement of the detector region is detected by reference to its size contribution to the overall amplification product of the tag and further primers. A convenient size difference may be used, even as little as one base pair difference can be detected on a gel. Generally however size differences of at least 5, conveniently at least 10, at 15 least 15 or at least 20 base pairs are used. This aspect of the invention is of particular use where two or more alleles of a genetic locus are to be detected in a single assay mixture.

The tag primer is capable of hybridisation to the complement of the tag sequence in the further primer extension product. It will be understood that the diagnostic primer extension product is separated from the further primer extension product prior to 20 hybridisation of the tag primer. The sequence of the tag primer is conveniently identical to the sequence of the tag region in the tail. The tag primer preferably comprises a sequence capable of hybridisation to all tag sequences. All tag sequences are preferably identical. Again it will be appreciated that minor changes may be made to the sequence of the tag primer without affecting its performance to any significant extent. The use of a common tag primer and 25 common tail sequences has significant cost advantages for a typical assay.

It will be understood that the diagnostic primer tail is non-complementary to any relevant genomic sequence or adjacent region so as not to compromise the assay.

In known diagnostic PCR procedures mispriming may occur at each amplification cycle, especially where the primer is used to detect for example single base mismatches or to 30 detect a particular sequence against a background of related sequences. Such mispriming may only occur as a very low percentage of total priming events per amplification cycle but will

increase significantly as a function of the overall number of cycles. The present invention comprises a two stage procedure wherein as a first stage the initial interaction between a diagnostic primer comprising tag and detector regions and a sample template may conducted at optimum hybridisation stringency. Any primer extension products are then amplified using 5 a further primer. As a second stage the above extension products are then amplified using a tag primer and the further primer. Accordingly, whilst mispriming may still initially occur the overall level may be significantly reduced.

As indicated above the tail sequences may be the same or different but are conveniently identical or substantially similar so that a single tail primer may be used. This 10 facilitates the performance of large multiplexes without overloading the reaction mix with different primers. We have found that the use of identical tag sequences can be advantageously used to even out the efficiencies of different amplification reactions.

We have also found that tailed primers can also be used to prevent the formation of "primer dimers" and other inter-primer artefacts. These are a particular problem in 15 homogeneous assays using for example intercalating dyes to detect double stranded nucleic acid. They result in false positive signals. See for example Ishiguro et al, Anal. Biochem., 1995, 229, 207-213, especially pages 211- 212. Whilst we do not wish to be limited by theoretical considerations, it is believed that the formation of primer dimers is dependent on some degree of homology between primers and their use at high concentrations. It may be 20 possible to reduce the formation of primer dimers by careful primer design. However where many primers are used at high concentrations, for example in PCR multiplexes, this becomes more difficult. We now disclose the use of diagnostic and further primers which are genome specific at their 3'-termini but which carry a detector region and common extensions (tags) at their 5'-termini. These are used in combination with a common tag primer which can prime 25 from the complement of the tag sequence in extension products of further primer(s). Thus whilst primer dimers and other inter-primer artefacts could occur during first phase diagnostic priming, these cannot be amplified during subsequent rounds of tag specific priming. It will be appreciated that the diagnostic primers are conveniently used at concentrations which allow satisfactory priming on their genomic template(s) but do not allow significant PCR 30 amplification.

The common tag primer is used at higher concentration than the genome specific primers.

To ensure that primer-dimers and other artefacts are avoided a common tag and common tag primer are preferably used for all primers present in a reaction mix,  
5 including control primers.

We have now found that it is advantageous to switch from diagnostic primer priming to tag primer priming by means of a temperature switch. The primers are selected so that the melting temperature of the tag primer is higher than the genome complementary region of the diagnostic primer. An increase in temperature will favour priming, for example after one or  
10 conveniently two rounds of diagnostic primer priming, by the tag primer.

The diagnostic primer may be an allele specific primer. In EP-A-0333465 (Baylor College of Medicine) there is described a detection method using two competing primers for the detection of diagnostic base sequences which differ by as little as a single base. This method depends on careful control of melting temperature (Tm) and is known as  
15 competitive oligonucleotide priming (COP). Competing primers may be used in the method of this invention, either the primers are differentially labelled or the amplification products are separated according to size, for example by the use of different size tails on the primers.

Furthermore in our European Patent, Publication No. 0332435, the contents of which are incorporated herein by reference, we disclose and claim a method for the selective  
20 amplification of template sequences which differ by as little as one base. The above method is now commonly referred to as the Amplification Refractory Mutation System (ARMS). This is of particular use, for example, where diagnostic base sequence(s) are only present in low concentration in complex nucleic acid mixtures.

Therefore in a preferred aspect of the above detection method a terminal  
25 nucleotide of at least one diagnostic primer is either complementary to a suspected variant nucleotide or to the corresponding normal nucleotide, such that an extension product of a diagnostic primer is synthesised when the terminal nucleotide of the diagnostic primer is complementary to the corresponding nucleotide in the diagnostic base sequence, no extension product being synthesised when the terminal nucleotide of the diagnostic primer is not  
30 complementary to the corresponding nucleotide in the diagnostic base sequence.

The diagnostic primers for use in the preceding aspect are conveniently designed with reference to our above mentioned European Patent, Publication No. 0332435.

By "substantially complementary" we mean that primer sequence need not reflect the exact sequence of the template provided that under hybridising conditions the primers are 5 capable of fulfilling their stated purpose. This applies equally to diagnostic and tail primers.

In general, mismatched bases are introduced into the primer sequence to provide altered extension rates with DNA polymerases. Commonly, however, the primers have exact complementarity except in so far as non-complementary nucleotides may be present at a predetermined primer terminus as hereinbefore described.

10           In the diagnosis of, for example, cancer the situation may arise whereby it is desirable to identify a small population of variant cells in a background of normal cells. The ARMS system is well suited for this purpose since it discriminates between normal and variant sequences even where the variant sequence comprises a very small fraction of the total DNA. Whilst we do not wish to be limited by theoretical considerations we have successfully 15 performed ARMS assays in which the ratio of mutant to normal DNA was 1:100 and we believe that even larger ratios may be readily used. To optimise the sensitivity of the ARMS reaction it may be performed in isolation ie. with a single ARMS primer since in duplex or multiplex reactions there may be competitive interaction between the individual reactions resulting in a loss of sensitivity. A control reaction is desirable to ensure that a polymerase 20 chain reaction has taken place. In a test for an inherited mutation the copy number of the mutation and other genomic is typically 1:1 or 1:2, so a genomic control reaction can be used without compromising sensitivity or creating an imbalance in the system. In a cancer test however, the use of a genomic control reaction may swamp the test reaction leading to a loss of sensitivity. We have now found that ARMS primer(s) comprising tail sequences may 25 advantageously be used in a two stage amplification procedure comprising a genomic control reaction. In the first stage ARMS primer(s) comprising non-complementary tail(s) are used to amplify any variant sequence which may be present. In addition to the ARMS reaction a genomic control reaction is performed in the same reaction vessel using primers at very low concentration. The control reaction primers also have non-homologous tails which may or 30 may not have the same sequence as the ARMS primer tail(s). In the second stage tail specific primers are added and the temperature increased to prevent the original genomic control

primers from functioning. In this second stage any variant sequence product is further amplified and the product of the control reaction from the first stage is also amplified to give a detectable product. Thus the ARMS reaction will only take place if variant sequence is present in the original sample and the control reaction will only function if both the first and 5 second stage amplification reactions have worked.

A further and important use of ARMS is for detecting the presence or absence of more than one suspected variant nucleotide in the same sample. The ability of ARMS to selectively amplify sequences depending on the predetermined nucleotide sequence of the diagnostic primers enables multiple amplification products to be distinguished simply, 10 accurately and with minimal operator skill thus making it possible to provide a robust technique for screening a single sample for multiple nucleotide variations. The use of ARMS to detect more than one suspected variant nucleotide in the same sample is conveniently referred to as multiplex ARMS. Multiplex ARMS is thus of particular interest in screening a single sample of DNA or RNA for a battery of inherited conditions such as genetic disorders, 15 predispositions and somatic mutations leading to various diseases. Such DNA or RNA may for example be extracted from blood or tissue material such as chorionic villi or amniotic cells by a variety of techniques such as those described by Maniatis *et al*, Molecular Cloning (1982), 280-281. Moreover as the molecular basis for further inherited conditions becomes known these further conditions may simply be included in the screening technique of the 20 present invention.

Multiple amplification products may be distinguished by a variety of techniques. Thus for example probes may be employed for each suspected amplified product, each probe carrying a different and distinguishable signal or residue capable of producing a signal.

25 A much simpler and preferred method of distinguishing between ARMS amplification products comprises selecting the nucleotide sequences of the amplification primers such that the length of each amplified product formed during the process of the present invention is different. In this regard the number of base pairs present in an amplification product is dictated by the distance apart of the diagnostic and amplification 30 primers. Thus the amplification primers may be designed such that each potential variant nucleotide is associated with a potential amplification product of different length.

In an ARMS reaction diagnostic for a particular point mutation the sequence of the primers is largely constrained by the sequence of the DNA adjacent the mutation of interest. The 3' base of the primer usually matches the base altered by the mutation and extra destabilisation is introduced to give the required level of specificity. The term "specificity" 5 refers to the ratio of the yield of product when an ARMS primer is used to prime its target sequence compared to the yield of mis-primed product from the non-target sequence.

In a multiplex ARMS reaction it is desirable that the individual ARMS reactions work with similar efficiency to allow the simultaneous detection of all the reaction products. This may be achieved for example by altering the concentration of the primers, 10 alteration of the number/composition of reactions, or alteration of the amount of additional destabilisations introduced into the ARMS primers. Whilst these methods are normally sufficient to obtain a balanced multiplex ARMS reaction the use of tail or tag sequences may have advantages in certain situations. In particular these may allow a more specific test. By way of example, where a strong additional mis-match is used to obtain specificity the yield of 15 corresponding multiplex product may be low. Reducing the additional mis-match strength may not be possible without compromising specificity. A tail sequence which in combination with a tail specific primer provides a good substrate for a DNA polymerase may be used to balance the multiplex reaction. A range of tail/primer combinations of known priming ability may be provided. Thus by way of example as a first amplification step the 20 priming/mis-priming ratio is optimised without regard to product yield. Product yield is then balanced in the second amplification step using an appropriate range of tail/primer combinations.

In our UK Patent No. 2252407 (Zeneca) we disclose and claim that multiplex ARMS may be successfully performed where diagnostic primer extension products of more 25 than one diagnostic base sequence of a nucleic acid sample comprise a complementary overlap. This unexpected improvement to multiplex ARMS is referred to hereinafter as overARMS. OverARMS now facilitates the detection and analysis of, for example, inherited or infectious disease where the potential variant nucleotides are closely spaced.

In an overARMS reaction the size of the reaction products can be used to identify 30 individual combinations of variant nucleotides. Where the products are separated for example on an agarose gel this approach may be limited by the resolving power of the gel. By way of

example in a high resolution agarose gel overARMS may presently be used to identify mutations within about 10-15 bases of each other. The size of the outer overARMS primer was increased to give a larger product and we surprisingly found that the yield of the smaller overARMS product was significantly reduced. Whilst we do not wish to be limited by 5 theoretical considerations we believe that target masking takes place due to the increased Tm of the larger overARMS primer which binds preferentially to the target DNA and prevents the smaller overARMS primer from hybridising. Use of a tailed outer overARMS primer may provide the increased product size necessary for resolution but since it is non-complementary at its 5' end the Tm will be similar to the smaller primer.

10 OverARMS is conveniently used for HLA typing, in the diagnosis of  $\beta$ -thalasaemia, sickle cell anaemia, phenylketonuria (PKU), Factor VIII and IX blood disorders and  $\alpha$ -1-antitrypsin deficiency. A particular use for OverARMS is in the detection and diagnosis of cystic fibrosis. Convenient cystic fibrosis alleles are disclosed in our European Patent Application No. 90309420.9; by B. Kerem *et al*, Science, 1989, 245, 1073-1080; by J.R. Riordan *et al*, Science, 1989, 245, 1066-1073; by J.M. Rommens *et al*, Science, 1989, 245, 1059-1065; by G.R. Cutting *et al*, Nature, 346, 366-368; by M. Dean *et al*, Cell, 61, 863-870; by K. Kobayashi *et al*, Am. J. Hum. Genet., 1990, 47, 611-615; by B. Kerem *et al*, Proc. Natl. Acad. Sci. USA, 1990, 87, 8447; by M. Vidaud *et al*, Human Genetics, 1990, 85, (4), 446-449; and by M.B. White *et al*, Nature, 344, 665-667.

15 Our two stage amplification process using diagnostic and tag primers in combination with a further common primer is conveniently carried out using all three primers simultaneously and preferably using a ratio of tail specific and/or further primer(s) to diagnostic primer(s) of at least 1:1, such as at least 20:1, at least 30:1, and at least 40:1, preferably at least 50:1.

20 All of the above detection methods involving PCR amplification may be provided as homogeneous assays.

The nucleic acid in the sample may be nucleic acid derived from for example viruses, bacteria (genomes and plasmids), bacteriophages, eukaryotic cells (nuclear, plasmid or organelle), humans, animals, plants, latent viruses in human or other cells. The sample is 25 conveniently obtained from an individual using conventional techniques. The nucleic acid

-10-

may be DNA, RNA or reverse transcribed RNA. It may be native, fragmented, cloned, degraded, extracted from cells or just released upon cell death.

An additional benefit of the method of the invention is in single tube genotyping ARMS assays. At a polymorphic locus in a diploid organism there are typically two 5 different alleles (A and B) and hence three possible genotypes (AA, AB and BB). One way to determine the genotype is to perform two separate ARMS reactions, one specific for allele A and the other specific for allele B. It is also possible to include both the A- and B-specific ARMS primers in a single reaction and use differential labelling or primer length to determine which of the primers have amplified. In practice there may be problems such as non- 10 specificity since the ARMS reaction product from one allele-specific primer may act as a target for mispriming for the other primer. However in the method of the present invention the initial extension reaction is from the tailed ARMS primer but subsequent amplification is via the tag primer. After, for example the second round of PCR the ARMS primer contributes very little to the amplification process and consequently the probability of inappropriate 15 priming from non-target reaction products is greatly reduced. The use of a temperature shift protocol after say two rounds of PCR to promote tag priming, will further reduce the chance of mis-priming. Detection of the products of different diagnostic primers is by differential product size or differential labelling.

The method of the invention may be used in combination with a number of known 20 detection systems. By way of example it may be used to improve the taqman assay as described for example by Holland et al, Proc.Natl. Acad. Sci.USA, 1991, 88, 7276-7280 and by Gelfand et al in US-A-5210015. A further assay is that described by Yamagata et al in EP-A-0 639647. A still further assay is the strand displacement assay (SDA), see for example Walker et al, Nucleic Acids Research, 1996, 24(2), 348-353, EP-A-0 678 581, EP-A-0 678 25 582, and EP-A-0 684 315.

The invention will now be further illustrated but not limited by reference to the following detailed description, Example, Table and Figures wherein:

Figures 1-4 show the effect of magnesium concentration on the fluorescence ratio of FAM/TAMRA in the TagMan embodiment of the invention.

-11-

Figure 1. The difference between Ya and the No DNA samples is strongest at 2.4 mM Mg. However, these conditions also favour ARMS mispriming and the difference between Ya and Yi is smaller; this is an ARMS dependent event.

Figures 2-4. Each of three replicates for Ya and No DNA is presented for three different Mg concentrations (Figure 2 - 1.2mM Mg, Figure 3 - 2.4mM Mg, Figure 4 - 4.4mM Mg). The reproducibility is good and the ratios between positive and negative samples is impressive.

Figure 5 shows a comparison of a TagMan assay of the invention with an optimised TaqMan assay, looking at an Insulin gene polymorphism). TagMan performed well. Indeed, at 2.4 mM Mg, the difference between positive and negative results was nearly two-fold better than that for the insulin amplicon. We believe that the TagMan probe anneals far more avidly than the amplifying TAG primer.

Figures 6-8 show comparisons with various controls. In order to be sure that signals obtained were produced as a result of the introduced reporter portion of the original primer, control experiments using tailed primers missing this portion were carried out (Figure 6 - 1.2 mM Mg, Figure 7 - 2.4 mM Mg, Figure 8 - 2.4 mM Mg). In all cases, the absence of the reporter region lead to fluorescence ratios comparable to the negative controls. No inappropriate probe cleavage was taking place under the conditions tested here.

Figure 9(a) shows genomic priming using the tailed three phase (3\*) primer of the invention. The tag region of the tail (abc) and the detector region (xyz) are shown as is the tag primer (abc).

Figure 9(b) shows complementary strand synthesis from the further primer (cba). A copy of the tailed 3\* primer has now been made.

Figure 10 shows the TagMan detection embodiment of the invention.

Figure 10(a) shows the further primer extension product (cba and arrow and following dotted line). The tag primer (abc) and probe (xyz) with attached fluorophore and quencher are shown and can now anneal to the copied tail

Figure 10(b) shows polymerase mediated extension of the taq primer, this encounters the hybridised probe and efficiently cleaves the probe, releasing the measured fluorophore away from its quencher. This is the same as conventional TaqMan.

Figure 10(c) shows continued amplification of the target region driven by the tag primer (abc) and efficient cleavage of the TaqMan probe (xyz). This allows real time or end point detection of the released fluorophore. The tag primer and the TaqMan probe are included in the PCR at high concentrations, while the long tailed primers are included at low concentrations, to maximise tag driven priming. In order to maximise the efficiency of the process, the TaqMan probe should anneal more strongly than the tag primer, otherwise cleavage will be inefficient. This can be achieved by manipulating the melting temperatures of the primers and their relative concentrations. When using TaqMan for allele discrimination the ASO element of the approach requires that the probe annealing is borderline to obtain maximum differentiation between the two variants. The new system is more easily optimisable because of the probe and drive elements are user-selected and can be optimised once for all amplicons.

Figure 11 shows the Molecular Beacons embodiment of the invention. Molecular beacons makes use of a similar quenching effect: at the ends of a probe are 5 bp sequences complementary to each other. At each end of the probe is a member of a pair of fluorophores, one absorbs the excitation light and emits it a wavelength which is quenched by the other fluorophore. At low temperatures, the complementary regions of the probe cause the formation of the stem loop structure, bringing the two fluorophores close to each other and amplified target, then it should hybridise, disrupting the hairpin and releasing the first fluorophore from the quenching effect of the second.

Figure 11(a) shows the further primer extension product (cba and arrow and following dotted line). The tag primer (abc) and molecular beacons probe (xyz) can now anneal to the copied tail. Continued amplification of the target region is driven by the tag primer, the middle portion of the primer is copied repeatedly.

Figure 11(b) shows a further primer extension product resulting from the continued amplification. The quenched beacons probe (xyz) hybridises to the copied middle section of the tag primer. It becomes stretched, releasing the emitting fluorophore from the quenching effect of the second fluorophore.

The Molecular Beacons method is a good way to detect amplicons but not so well suited for allele discrimination. In our new scheme, allele discrimination can be obtained via ARMS but more importantly, the probe region may be designed with the idea of maximising

the stretching out effect caused by the hybridisation. In particular, the last 5 bp on the ends of the probe can be made to hybridise to the target rather than "flapping around"; this should allow the use of shorter probes and yield stronger signals. There is no need to design and produce new and expensive probes for each new amplicon, a generic probe can be used for 5 many different amplification targets. The system is multiplexible by changing the probe and introduced target sequences, and using a different pair of matched fluorophores.

Figure 12 shows the FRET detection embodiment of the invention. The basic method for this approach is to introduce two probes each carrying a member of a fluorophore pair. When the two probes hybridise to their amplified targets (which are essentially adjacent 10 to each other), the absorbed energy from exciting the first fluorophore is transferred to the second fluorophore which then emits at its characteristic wavelength. This can be greatly shifted from the excitation and emission wavelengths of the first fluorophore and produces very low backgrounds. In this format, the spacing between the two probes is crucial and requires case-by-case optimisation. Furthermore, the use of two fluorescent probes for each 15 amplicon is an expensive and cumbersome path.

Figure 12(a) shows continued amplification of the target region driven by the tag primer (abc) in combination with the further primer (cba). The middle portion of the primer (detector region) is copied repeatedly. Also shown are two probes (x,y; and z) which each carry half of an energy transfer pair.

20 Figure 12(a) shows how, after amplification, the two probes hybridise to the copied middle section of the tag primer. This allows energy transfer between the probes thus generating a detectable signal, with low background.

By allowing the user to define the probe sites, a single probe pair can be designed and optimised. This is then suitable for use against any target. This system can be multiplexed by 25 simply changing the middle portion of the original primer and using a different pair of fluorophores on appropriate probes. A refinement of this technique, suitable for real-time assays, is to use directly abutting probes which, once hybridised to the introduced target can be ligated together, for example using a thermostable ligase, and thus fixed in their fluorescing configuration. The ligated double probe can then be displaced by the (taq) 30 polymerase. The ligated product may be modified to prevent (TaqMan) cleavage of the newly adjoined sequences.

Figure 13 shows the Capture and Detection embodiment of the invention. The introduced middle section which becomes copied can be used as the sequence by which specific PCR products are captured. Post capture, a detector molecule can be introduced by way of a second target related probe and detection proceeds as appropriate. In this way, 5 mutant and normal amplicons can be decoded, simply by having unique capture sequence.

Figure 13(a) shows continued amplification of the target region driven by the tag primer (abc) in combination with the further primer (cba). The middle section of the further primer extension product is copied repeatedly. Each amplicon may have a characteristic capture signature (ie. xyz can differ).

10 Figure 13(b) shows how the middle portion (x'y'z') of the further primer extension product is a target for capture by an immobilised probe (xyz). A further probe (lmn) which carries a label (eg. biotin, horse raddish peroxidase, or alkaline phosphatase) may be hybridised to the amplicon for subsequent detection. In this way mutant and normal sequences may be detected in a single vessel, such as a tube, by using different capture 15 regions to discriminate between the two products.

Figure 14 shows the Lanthanide Enhanced Genetics System (LEGS) embodiment of the invention. LEGS is disclosed in our PCT patent application no. WO-95/08642. In LEGS, a partially caged Europium ion is attached to a probe. When a PCR product is rendered single stranded and hybridised to such a probe, a double-stranded region is produced. Also present 20 in the mixture is a synthetic Intercalator/Sensitiser (I/S) molecule which intercalates in the double-stranded region. The intercalator also has a linker arm which ends with a second caging group. In an appropriate conformation, the partially caged Eu ion can become fully caged thanks to the I/S molecule. This excludes and leads to a strong, time-resolved fluorescence effect. The I/S molecule is difficult to synthesise and may not be fully heat 25 stable.

Using the three phase primer of this invention, the middle phase can be constructed to allow a probe carrying a chelate of Eu to hybridise adjacent to a second probe bearing another caging group. This generates a fully caged Eu ion which then fluoresces. A major advantage of this system is that no complex organic synthesis is required. In addition, the second 30 chelator can be introduced in a targeted manner. Using known techniques, both the Eu chelated probe and the cage probe can be readily synthesised.

Figure 14(a) shows continued amplification of the target region driven by the tag primer (abc) in combination with the further primer (cba). The middle section of the further primer extension product is copied repeatedly. The mixture also contains two probes (x,y: and z), one of which carries a chelated lanthanide, the other carries a second chelating group.

5 Figure 14(b) shows how, after amplification, the two probes hybridise to the copied middle section of the further primer extension product. This causes complete caging of the lanthanide ion resulting in high efficiency time resolved fluorescence.

Figure 15 shows probe cleavage detection methods other than using the 5' exonuclease activity of taq polymerase. The probe and the introduced target may differ by 10 one or more bases, rendering any duplex formed between the two susceptible to cleavage by a number of methods, such as chemical cleavage or "cleavase" enzyme. Other approaches include the introduction of a restriction site on the middle position of the primer. After PCR, this may be cleaved releasing the detected fluorophore from its quencher. The newly synthesised DNA may be restriction endonuclease resistant (e.g. by using 15 methylated or phosphothioate dNTPs). We then allow the unprotected probes to be nicked when annealed to the target. Using a thermostable endonuclease would render this assay format fully homogeneous and suitable for real time detection. Alternatively, some enzymes require methylated double-stranded DNA for cleavage and these in combination with a methylated probe provide a further detection system.

20 Figure 15(a) shows shows the further primer extension product (cba and arrow and following dotted line). The tag primer (abc) and probe (xyz) with attached fluorophore and quencher are shown. These may anneal to the copied tail.

Figure 15(b) shows the probe (xyz) being cleaved using chemical or enzymatic methods.

25 Figure 16 shows other binding based methods. The introduced and copied, middle segment of the primer may also be designed as the target for a number of specific binding techniques, which may be suitable for product detection. For example, a triple helix motif may be introduced allowing specific detection of completed double-stranded amplicons. Another example is mismatched probes detected by mismatch binding proteins. Other 30 sequence specific protein binding events may be suitable for detection of the amplified middle

-16-

segment. Finally, simple FP probes may be detected and enhanced by having protein binding superimposed upon the probe/target hybridisation.

Figure 16(a) shows continued amplification of the target region driven by the tag primer (abc) in combination with the further primer (cba). The middle section of the further 5 primer extension product is copied repeatedly. Each amplicon may have a characteristic capture signature (ie. xyz can differ).

Figure 16(b) shows how the copy of the middle portion (x'y'z') of the further primer extension product is a target for subsequent binding of a number of detector molecules. A triple helix is formed using probe x\*y\*z\*.

10 Figure 17 shows single tube genotyping using differentially labelled 3\* ARMS primers.

Figure 17(a) shows allele specific extension of differentially labelled ARMS 3\* primers (abcxyz and abcrst) on targets sequences (alleles A and B). Also shown is a common tag primer (abc).

15 Figure 17(b) shows that subsequent amplification for both allele A and B specific ARMS products is from the common tag primer (abc). Products from the A specific ARMS primer now contain the reporter sequence xyz, whereas products from the B specific ARMS primer contain the reporter sequence rst. Differential detection is conveniently by size (xyz and rst are of different lengths) or by signal detection (xyz and rst reporter groups produce 20 different signals).

#### Example 1

#### Materials and Methods

##### **Primers and Probes**

25

Code	Comment	(5') Amplifier Portion	(Middle) Reporter portion	(3') Genome Priming portion
S7970	Amplifies Tagged amplicon	5'CGTACCAACGT GTGCACT3'	NONE	NONE

-17-

HpH1F	insulin gene	5'AGCAGGTCTG TTCCAAGG3'	NONE	Primer/amplifier sequence
HpH1R	insulin gene	5'CTTGGGTGTG TAGAAGAAC3'	NONE	Primer/amplifier sequence
INS-BF1	insulin gene-probe	NONE	5'FAM-CCTGCCTGTCT CCCAGATCAC TAMRA3'	NONE
S7033	ΔF508 common primer (Amplifier tailed)	5'CGTACCACGT GTCGACT3'	NONE	5'CACTAATG AGTGAACAA AATTCTCACC ATT3'
T2120	TagMan Probe	NONE	5'FAM-CTGGCATCGG TAGGGTAAGG ATCGGTATCGT AMRA3'	NONE
T0990	Triple-phase primer for ΔF508 amplicon	5'CGTACCACGT GTCGACT3'	5'(GCGTACT)C TGGCATCGGT AGGGTAAGGA TCGGTATCG3'	5'GCCTGGCA CCATTAAAG AAAATATCA TTGG3' Bracketed portion is an optional “hinge” region

P1292	No reporter target primer	5'(GCGTACT)CG TACCACGTGTC GAAT3' Bracketed portion is an optional "hinge" region	NONE	5'GCCTGGCA CCATTAAAG AAAATATCA TTGG3'
-------	------------------------------	-------------------------------------------------------------------------------------------------	------	------------------------------------------------

## PCRs

1. Insulin amplicon: 25 µl reactions containing
  - 5 3mM MgCl<sub>2</sub>
  - 200 µM dNTPs
  - 10% (v/v) glycerol
  - 3 ng/ml each of primers HpH1F and HpH1R
  - 50 nM INS-B1F probe
  - 10 5 µl genomic DNA (or water for negative controls)
  - 0.625 units of Taq polymerase
  - in 1x AmpliTaq Buffer
- 15 These reactions were performed using a two step cycle:  
40 cycles x {94 °C for 1 min; 58 °C for 2 min}
2. Ya reactions (150 µl) contained 1 x ARMS buffer with: 100 µM dNTPs. Primers T0990 and S7033 at 10 nM, S7970 at 500 nM, T2120 at 50 mM, 225 ng ΔF508 homozygote DNA and 6 unites of Taq Polymerase. Some reactions were supplemented with Mg to 2.4 or 4.4 mM
- 20 Yi reactions were identical to Ya, but the target DNA was a normal homozygote at the F508 position.
- Other controls contained no target DNA

To control for the reporter sequence in the Triple phase primer, P1292 (containing no middle portion) was substituted for T0900.

The PCR cycles were: 2x{94°C for 1 min, 62°C for 2 min, 72°C for 1 min}, followed by 40x {94°C for 1 min, 64°C for 1 min}.

5

### Analysis

After cycling, an aliquot (10 µl) of each reaction was analysed by gel electrophoresis to establish the efficiency of amplification.

The remainder was analysed in 100 µl cuvettes using the Fluoromax fluorometer. Where 10 necessary (as in the case of the insulin amplifications), replicate samples were pooled. The excitation wavelength was set to 488 nm and the emission was read at 518 nm (for FAM) and 582 nm (for TAMRA). The ratios were calculated in each case and plotted appropriately.

### Results

15 The optimal [Mg] for the insulin gene region had been established as 3 mM (John Todd pers. comm.) Indeed, there was little or no change in the fluoresence ratio at lower concentrations. This was borne out in the data shown in Figure 1. The difference between Ya and the No DNA samples is strongest at 2.4 mM Mg. However, these conditions also favour ARMS mispriming and the difference between Ya and Yi is less impressive: this is due to 20 ARMS not TagMan.

Each of three replicates for Ya and no DNA is presented (Figures 2-4) for three different Mg concentrations. The reproducibility is good and the ratios between positive and negative samples in impressive.

When compared to the results obtained in an optimised TaqMan assay (the Insulin 25 gene polymorphism) the method of this invention performed well (Figure 5). Indeed, at 2.4 mM Mg, the difference between positive and negative results was nearly two-fold better than that for the insulin amplicon. This probably reflects the fact that we were able to design the TagMan probe to anneal far more avidly than the amplifying tag primer.

In order to be sure that signals obtained were produced as a result of the introduced 30 reporter portion of the original primer, control experiments using tailed primers missing this portion were carried out (Figures 6-8). In all cases, the absence of the reporter region lead to

-20-

fluorescence ratios comparable to the negative controls. No inappropriate probe cleavage was taking place under the conditions tested here.

To demonstrate the universality of this approach, a number of further amplicons have been  
5 worked up to 3\* format. In each case, the separate allele specific reactions have been successfully combined to permit Single Tube Genotyping (STG). In one case (factor V Leiden) the assay has been blind tested extensively on clinical samples (which were typed independently by a clinical laboratory using a different technology).

#### 10 1. Cystic Fibrosis Delta F508

Primers (see Table 1):

- V6631: Tag 20a  
V6632 common primer, tailed with 20a sequence  
V6634: 3\* primer, mutant sequence, T2120 reporter  
15 V6634: 3\* primer, normal sequence, T4029 reporter

Probes:

- T2120: FAM/TAMRA labelled,  
T4029: TET/TAMRA labelled,

Reaction mixes:

- 20 All reactions were in 1xARMS buffer with MgCl<sub>2</sub> adjusted to 3.5 mM final, plus ROX internal standard at 60 nM final, and Tag20a at 500 nM. AmpliTaq Gold was included at 2 U per 50 µl reaction

Three mixes were typically used-

- a. Normal only reaction, with primers V6635 and V6632 each at 10 nM, T4029 at  
25 50 nM  
b. Mutant only reaction, with primers V6634 and V6632 each at 10 nM, T2120 at 50nM  
c. STG with all three primers included at 10 nM, both probes at 50 nM each

Cycling conditions:

- 30 20 minutes at 94°C to activate the AmpliTaq Gold  
2 cycles of 94°C, 40s; 62°C, 80s; 72°C, 40s (genomic priming)

-21-

40 cycles of 94°C, 4s, 62°C, 80s (Tag priming)

## 2. CTLA4A polymorphism

Primers (see Table 1):

- 5 V6631: Tag20a
- V3561: Common primer, tailed with Tag 20a sequence
- V6558: 3\* primer, mutant specific, carries the TET reporter region (T4029)
- V6715: 3\* primer, normal specific, carries the FAM reporter region (T2120)

Probes:

- 10 T2120: FAM/TAMRA labelled,
- T4029: TET/TAMRA labelled,

Reaction conditions:

All reactions were in 1xARMS buffer with MgCl<sub>2</sub> adjusted to 3.5 mM final, plus ROX internal standard at 60 nM final, and Tag20a at 500 nM. AmpliTaq Gold was included at 2

- 15 U per 50 µl reaction

Three mixes were typically used-

- a. Normal only reaction, with primers V6715 and V3561 each at 10 nM, T2120 at 50 nM
- b. Mutant only reaction, with primers V6558 and V3561 each at 10 nM, T 4029 at 50 nM
- c. STG with all three primers included at 10 nM, both probes at 50 nM each

Cycling conditions:

20 minutes at 94°C to activate the AmpliTaq Gold

3 cycles of 94°C, 40s; 64°C, 80s; 72°C, 40s (genomic priming)

- 25 40 cycles of 94°C, 4s, 64°C, 80s (Tag priming)

## 3. BRCA2exon10 polymorphism

Primers (see Table 1):

- V6631: Tag 20a
- 30 R432-96: common primer, tailed with Tag 20 sequence
- V9596: 3\* primer, A-variant specific, carries T4029 (TET) reporter sequence

-22-

W1940: 3\* primer, C-variant specific, carries T2120 (FAM) reporter sequence

Probes

T2120: FAM/TAMRA labelled,

T4029: TET/TAMRA labelled,

5 Reaction Conditions:

All reactions were in 1xARMS buffer with MgCl<sub>2</sub> adjusted to 3.5 mM final, plus ROX internal standard at 60 nM final, and Tag20a at 500 nM. AmpliTaq Gold was included at 2 U per 50 µl reaction

Three mixes were typically used-

- 10 a. "A" only reaction, with primers R432-96 and V9596 each at 25 nM, T4029 at 50 nM  
b. "G" only reaction, with primers R432-96 and W1940 each at 25 nM, T2120 at 50 nM  
c. STG with all three primers included at 25 nM, both probes at 50 nM

15 Cycling conditions:

20 minutes at 94°C to activate the AmpliTaq Gold

4 cycles of 94°C, 40s; 60°C, 80s; 72°C, 40s (genomic priming)

45 cycles of 94°C, 40s, 64°C, 80s (Tag priming)

20 4. Factor V Leiden mutation

Primers (see Table 1):

V0651: 3\* primer (wild type sequence) with reporter region corresponding to

T2120

V0652: 3\* primer (mutant sequence) with reporter region corresponding to T4029

25 W4085: Tailed common primer

Tag 20a: Driver primer found in the common and the specific primers

Probes:

T2120: FAM/TAMRA labelled,

T4029: TET/TAMRA labelled,

30 Reaction mixes:

All reactions were in 1xARMS buffer with MgCl<sub>2</sub> adjusted to 3.5 mM final, plus ROX

-23-

internal standard at 60 nM final, and Tag20a at 500 nM. AmpliTaq Gold was included at 2 U per 50 µl reaction

Three mixes were typically used-

- 5    a.    Normal only reaction, with primers W4085 and V0651 each at 25 nM, T2120 at 50 nM
- b.    Mutant only reaction, with primers W4085 and V0652 each at 25 nM, T4029 at 50 nM
- c.    STG with all three primers included at 25 nM, both probes at 50 nM

#### 10 Cycling conditions:

20 minutes at 94°C to activate the AmpliTaq Gold

3 cycles of    94°C, 41s; 60°C, 80s; 72°C, 51s (genomic priming)

45 cycles of    94°C, 41s, 66°C, 80s (Tag priming)

#### 15 Validation

More than 200 clinical samples have been tested blind with the STG mix. In every case, the results obtained were concordant with those obtained by clinical collaborators who used PCR and restriction digestion to type the same samples.

**TABLE 1**

<b><u>CODE</u></b>	<b><u>SEQUENCE</u></b>
V6631	GCGTACTAGCGTACCAACGTG
T4029	CGGTGGACGTGACGGTACGACGAGGCGACG
T2120	CTGGCATCGGTAGGGTAAGGATCGGTATCG
V6632	GCGTACTAGCGTACCAACGTGCACTAATGAGTGAACAAAATTCTCA CCATT
V6635	GCGTACTAGCGTACCAACGTGTCGACTCGGTGGACGTGACGGTACG ACGAGGCGACGGCCTGGCACCAATTAAAGAAAATATCATCTT
V6634	GCGTACTAGCGTACCAACGTGTCGACTCTGGCATCGTAGGGTAAG GATCGGTATCGGCCTGGCACCAATTAAAGAAAATATCATTGG
V3561	GCGTACTAGCGTACCAACGTGATCCTGAAACCCAGCTCAAAT
V6558	GCGTACTAGCGTACCAACGTGTCGACTCGGTGGACGTGACGGTACG ACGAGGCGACGGCGGCACAAATAAAACTGAACCTGGCTG
V6715	GCGTACTAGCGTACCAACGTGTCGACTCTGGCATCGTAGGGTAAG GATCGGTATCGGCAGCACAAATAAAACTGAACCTGGCTA
V0652	GCGTACTAGCGTACCAACGTGTCGACTCGGTGGACGTGACGGTACG ACGAGGCGACGTACTCAAGGACAAATACCTGTATTCCAT
V0651	GCGTACTAGCGTACCAACGTGTCGACTCTGGCATCGTAGGGTAAG GATCGGTATCGTACTCAAGGACAAATACCTGTATTCCGC

W4085      GCGTACTAGCGTACCAACGTGCAGGGAAACCTATACTTATAAGTG  
              GAACATC

R432-96     GCGTACTAGCGTACCAACGTGAGAAGTTCCAGATATTGCCTGCTT

V9596      GCGTACTAGCGTACCAACGTGTCGACTCGGTGGACGTGACGGTACG  
              ACGAGGCGACGACTGATCCATTAGATTCAAATGTAGGAA

W1940      GCGTACTAGCGTACCAACGTGTCGACTCTGGCATCGGTAGGGTAAG  
              GATCGGTATCGACTGATCCATTAGATTCAAATGTAGAAC

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

5

## (i) APPLICANT:

- (A) NAME: Zeneca Limited
  - (B) STREET: 15 Stanhope Gate
  - (C) CITY: London
- 10 (E) COUNTRY: United Kingdom
- (F) POSTAL CODE (ZIP): W1Y 6LN
  - (G) TELEPHONE: 0171 304 5000
  - (H) TELEFAX: 0171 304 5151
  - (I) TELEX: 0171 834 2042

15

## (ii) TITLE OF INVENTION: Process

## (iii) NUMBER OF SEQUENCES: 15

## 20 (iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

25

## (v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: GB 9609441.2

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

- 5       (A) LENGTH: 20 base pairs  
      (B) TYPE: nucleic acid  
      (C) STRANDEDNESS: single  
      (D) TOPOLOGY: linear

10     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GCGTACTAGC GTACCACGTG

20

## (2) INFORMATION FOR SEQ ID NO: 2:

15

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs  
      (B) TYPE: nucleic acid  
      (C) STRANDEDNESS: single  
20     (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

25 CGGTGGACGT GACGGTACGA CGAGGCGACG

30

-28-

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- 5 (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

10

CTGGCATCGG TAGGGTAAGG ATCGGTATCG

30

(2) INFORMATION FOR SEQ ID NO: 4:

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GCGTACTAGC GTACCACGTG CACTAATGAG TGAACAAAAT TCTCACCATT

50

25

30

-29-

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 86 base pairs
- 5 (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GCGTACTAGC GTACCACGTG TCGACTCGGT GGACGTGACG GTACGACGAG  
GCGACGGCCT 60

15 GGCACCATTA AAGAAAATAT CATCTT

86

(2) INFORMATION FOR SEQ ID NO: 6:

20 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 86 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GCGTACTAGC GTACCACGTG TCGACTCTGG CATCGTAGG GTAAGGATCG  
GTATCGGCCT 60

30

GGCACCATTA AAGAAAATAT CATTGG

86

-30-

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- 5 (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

10

GCGTACTAGC GTACCACGTG ATCCTGAAAC CCAGCTAAA T

41

(2) INFORMATION FOR SEQ ID NO: 8:

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 85 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GCGTACTAGC GTACCACGTG TCGACTCGGT GGACGTGACG GTACGACGAG

GCGACGGCGG 60

25

CACAAATAAA AACTGAACCT GGCTG

85

30

## (2) INFORMATION FOR SEQ ID NO: 9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 85 base pairs
- 5 (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

10

GCGTACTAGC GTACCACGTG TCGACTCTGG CATCGTAGG GTAAGGATCG  
GTATCGGC GG 60

CACAAATAAA AACTGAACCT GGCTA

85

15

## (2) INFORMATION FOR SEQ ID NO: 10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 86 base pairs
- 20 (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

25

GCGTACTAGC GTACCACGTG TCGACTCGGT GGACGTGACCG GTACGACGAG  
GCGACGTACT 60

TCAAGGACAA AATACCTGTA TTCCAT

86

30

-32-

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 86 base pairs
- 5 (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

10

GCGTACTAGC GTACCACGTG TCGACTCTGG CATCGTAGG GTAAGGATCG

GTATCGTACT 60

TCAAGGACAA AATACCTGTA TTCCGC

86

15

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 52 base pairs
- 20 (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

25

GCGTACTAGC GTACCACGTG CAGGGGAAAC CTATACTTAT AAGTGGAACA TC

52

30

-33-

(2) INFORMATION FOR SEQ ID NO: 13:

5 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GCGTACTAGC GTACCACGTG AGAACGTTCCA GATATTGCCT GCTT

44

(2) INFORMATION FOR SEQ ID NO: 14:

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 84 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GCGTACTAGC GTACCACGTG TCGACTCGGT GGACGTGACG GTACGACGAG

25 GCGACGACTG 60

ATCCATTAGA TTCAAATGTA GGAA

84

30

-34-

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 84 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GCGTACTAGC GTACCACGTG TCGACTCTGG CATCGGTAGG GTAAGGATCG  
GTATCGACTG 60

15 ATCCATTAGA TTCAAATGTA GAAC

84

**Claims:**

1. A method for the detection of a diagnostic base sequence in nucleic acid comprised  
5 in a sample, which method comprises contacting the sample under hybridising conditions and  
in the presence of appropriate nucleoside triphosphates and an agent for polymerisation  
thereof, with a diagnostic primer for the diagnostic base sequence, the diagnostic primer  
having a non-complementary tail sequence comprising a tag region and a detector region,  
such that an extension product of the diagnostic primer is synthesised when the corresponding  
10 diagnostic base sequence is present in the sample, no extension product being synthesised  
when the corresponding diagnostic base sequence is not present in the sample and any  
extension product of the diagnostic primer acts as template for extension of a further primer  
which hybridises to a locus at a distance from the diagnostic base sequence, and contacting  
the sample with a tag primer which selectively hybridises to the complement of the tag  
15 sequence in an extension product of the further primer and is extended, and detecting the  
presence or absence of the diagnostic base sequence by reference to the detector region in the  
further primer extension product.
2. A method as claimed in claim 1 wherein a detector species is used which selectively  
20 associates with the detector region in the further primer extension product.
3. A method as claimed in claim 2 wherein the detector species emits a detectable  
signal when cleaved during polymerase mediated extension of the tag primer.
- 25 4. A method as claimed in claim 2 wherein the detector species emits a detectable  
signal upon selective association with the detector region.
5. A method as claimed in claim 4 wherein the detector species is a fluorescently  
labelled species and the detectable signal arises from a change in fluorescence polarisation  
30 upon selective association with the detector region.

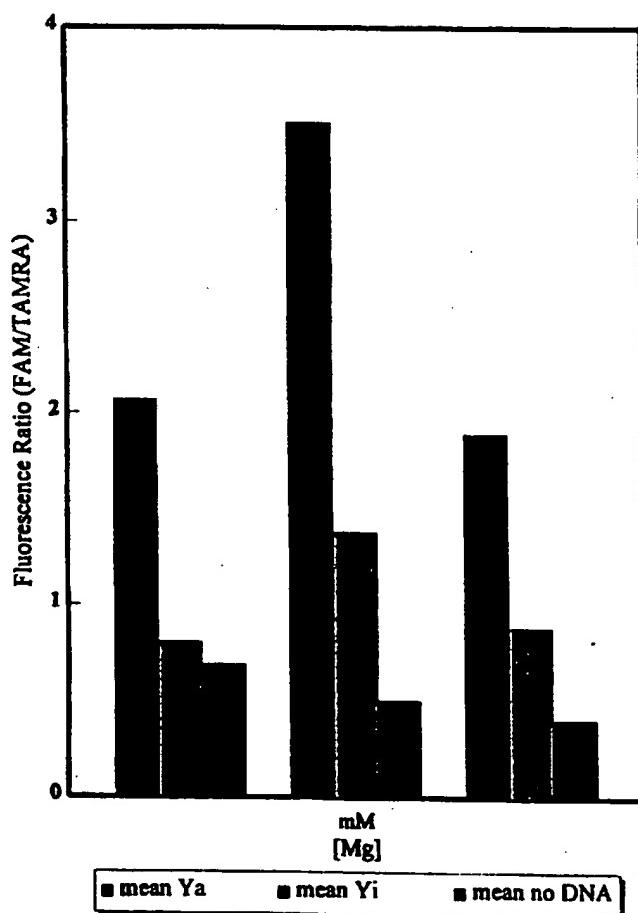
6. A method as claimed in claim 2 wherein the detector species comprises two species each having an interactive label, which labels interact upon selective association with the detector region and release a detectable signal.
- 5 7. A method as claimed in claim 6 wherein one of the interactive labels is a chelated lanthanide and the other is a further chelating group.
8. A method as claimed in claim 1 wherein the further primer extension product is captured on a solid phase using a species which selectively associates with the complement of 10 the detector region in the further primer extension product.
9. A method as claimed in any one of the previous claims wherein the detector species comprises a nucleotide sequence identical to the sequence of the detector region in the tail of the diagnostic primer.
- 15
10. A method as claimed in any one of the previous claims wherein the tag primer comprises a nucleotide sequence identical to the sequence of the tag region in the tail of the diagnostic primer.
- 20 11. A method as claimed in claim 1 wherein the detector region in the further primer extension product is identified by reference to its size contribution to the further primer/tag primer amplification product.
12. A method as claimed in any one of the previous claims and wherein the further 25 primer has a non-complementary tail sequence comprising a tag region.
13. A method as claimed in any one of the previous claims and wherein the further primer is a diagnostic primer.

14. A method as claimed in any one of the previous claims wherein the melting temperature of the tag primer is higher than that of the diagnostic primer so that an increase in temperature provides a switch from diagnostic primer priming to tag primer priming.
- 5 15. A method as claimed in any one of the previous claims wherein more than one diagnostic base sequence is detected in the sample using more than one diagnostic primer, appropriate further primer(s) and tag regions.
16. A method as claimed in claim 15 wherein the same tag sequence is used in the tail of 10 all diagnostic primers and/or all further primers.
17. A method as claimed in claim 15 wherein an identical detector region is used in the tail of all diagnostic primers.
- 15 18. A method as claimed in claim 15 wherein one further primer is used with more than one diagnostic primer.
19. A method as claimed in any one of the previous claims wherein a terminal nucleotide of at least one diagnostic primer is either complementary to a suspected variant 20 nucleotide or to the corresponding normal nucleotide.
20. A diagnostic primer having a non-complementary tail sequence comprising a tag region and a detector region.
- 25 21. A diagnostic primer having a 3' terminal nucleotide either complementary to a suspected variant nucleotide or to the corresponding normal nucleotide and having a non-complementary tail sequence comprising a tag region and a detector region.
22. A method for the identification of one or more variant diagnostic base sequences 30 against a background of normal diagnostic base sequences which comprised the use of a method as claimed in claims 1-19 or a diagnostic primer as claimed in claims 20-21.

-38-

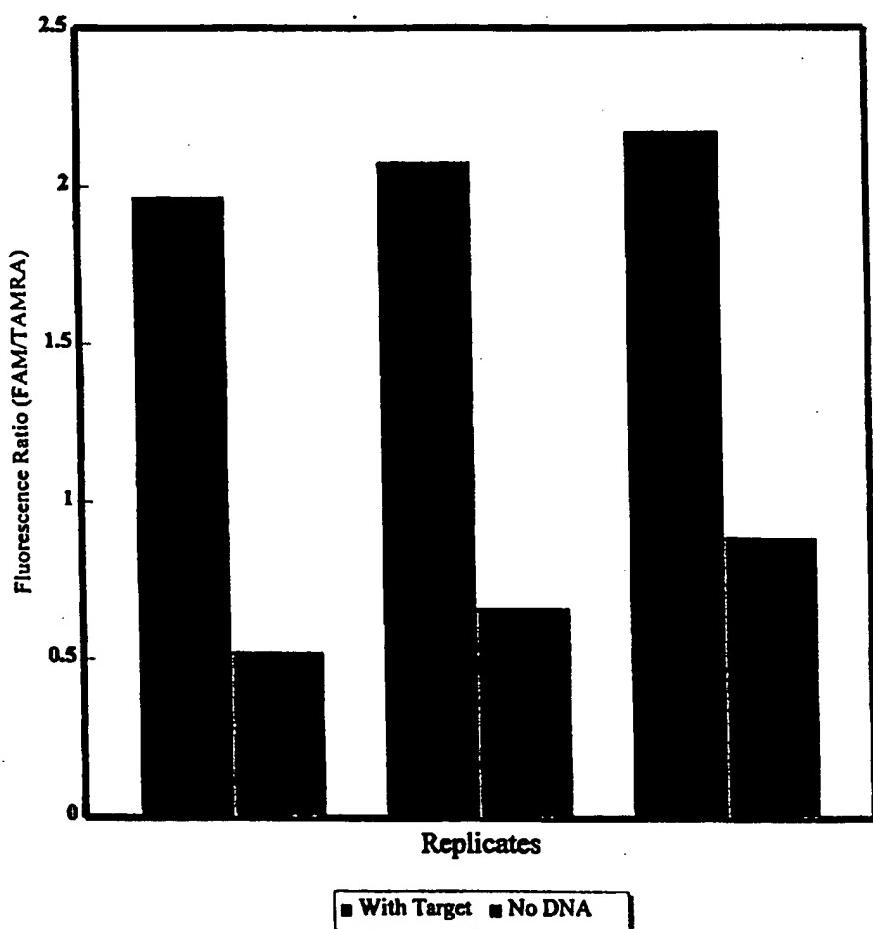
23. A kit comprising one or more diagnostic primers as claimed in claim 20 or claim 21 together with appropriate packaging and instructions for use in a method as claimed in any one of claims 1-19 or 22.

1/17

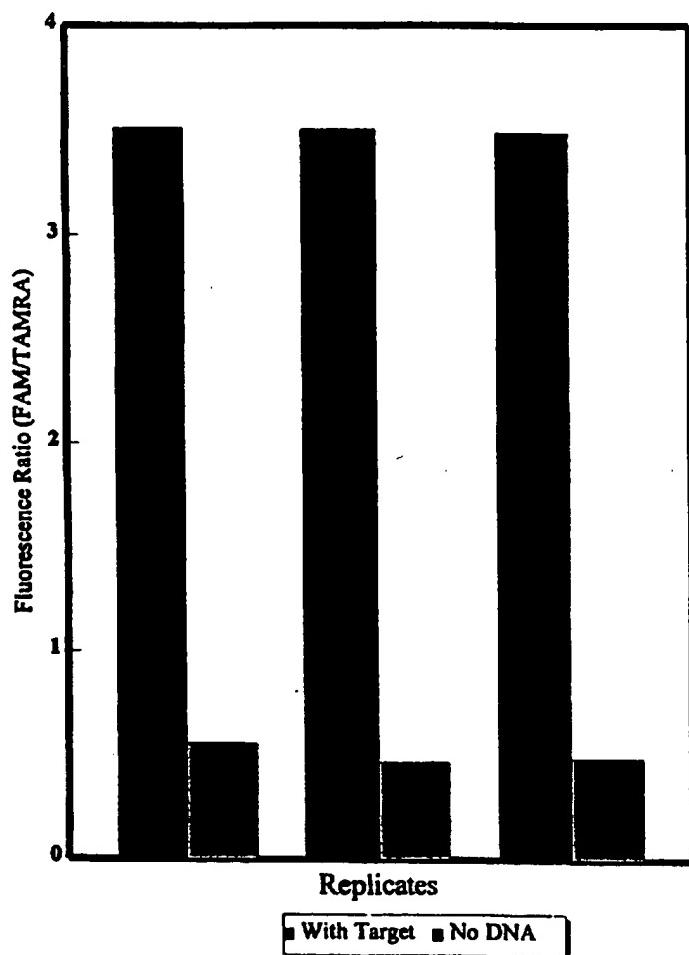
**FIGURE 1**

SUBSTITUTE SHEET (RULE 26)

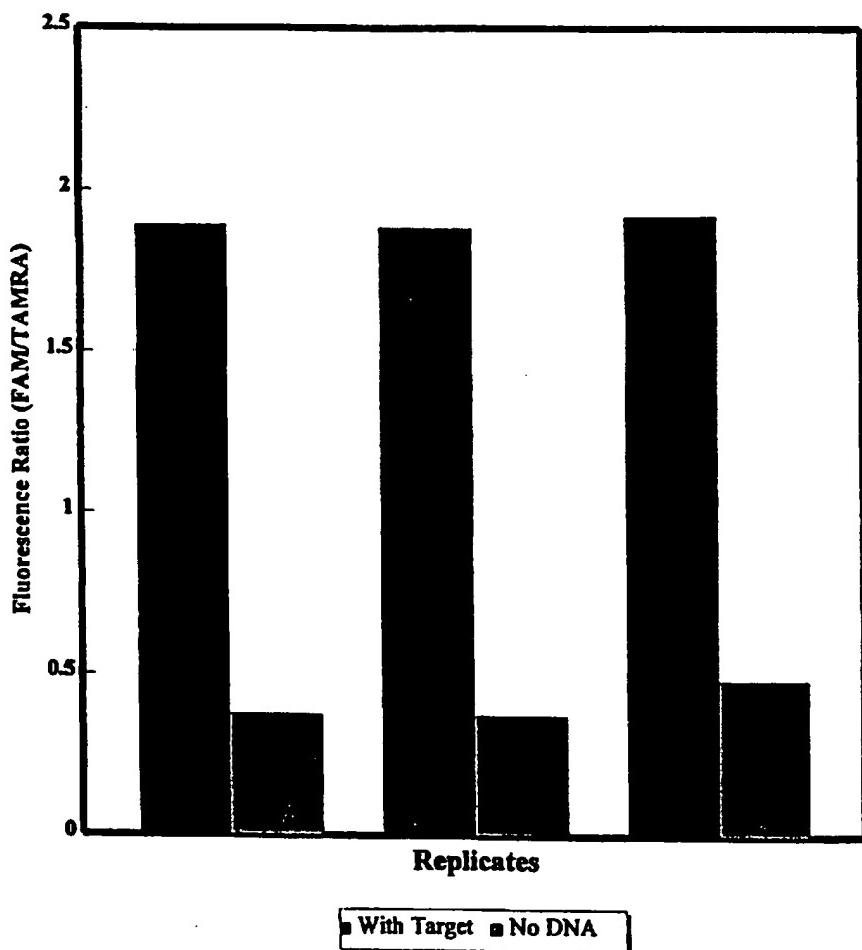
2/17

**FIGURE 2****SUBSTITUTE SHEET (RULE 26)**

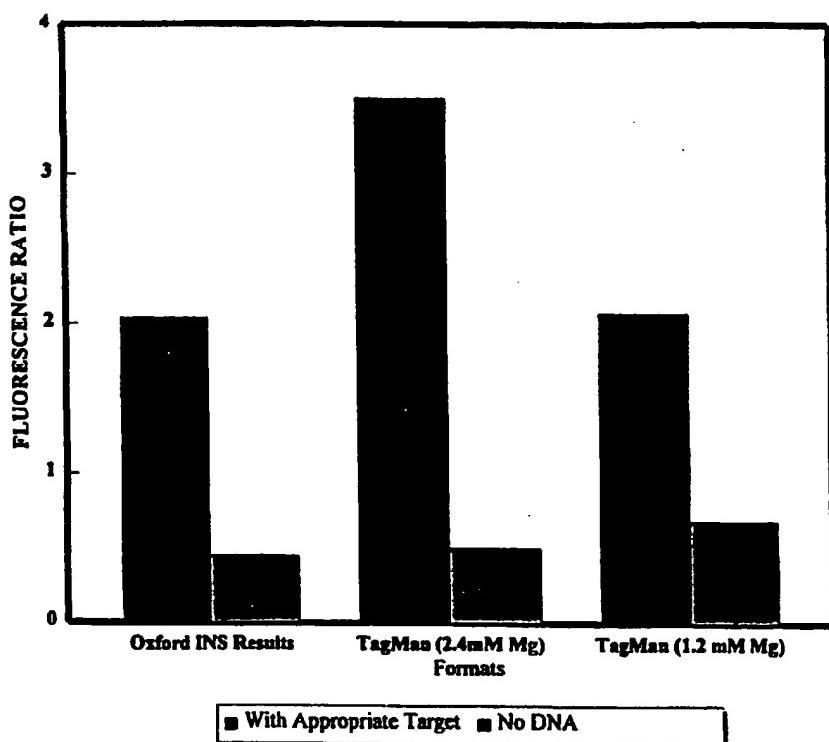
3/17

**FIGURE 3****SUBSTITUTE SHEET (RULE 26)**

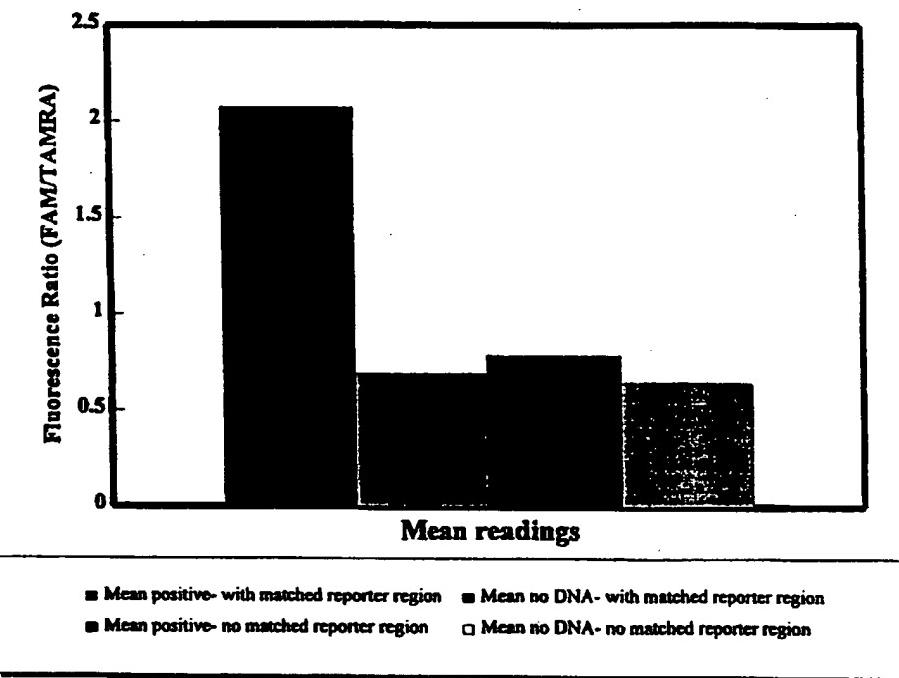
4/17

**FIGURE 4**

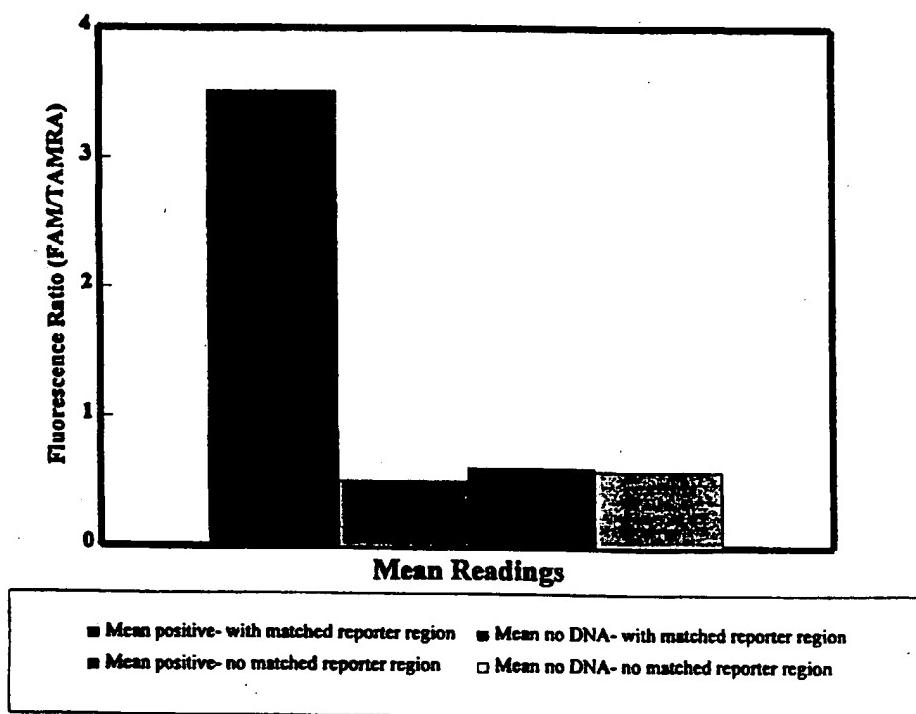
5/17

**FIGURE 5****SUBSTITUTE SHEET (RULE 26)**

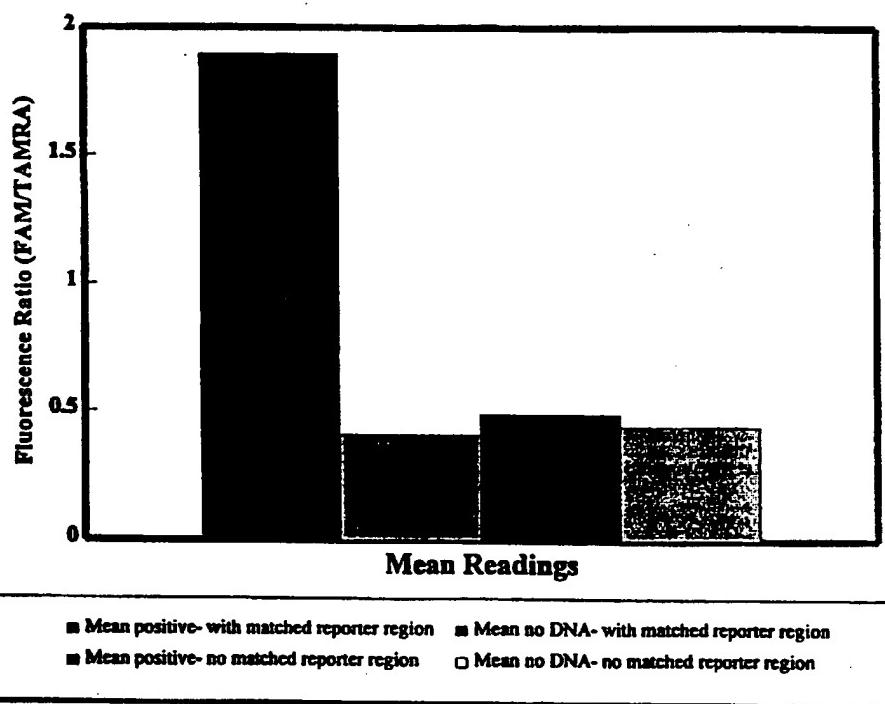
6/17

**FIGURE 6****SUBSTITUTE SHEET (RULE 26)**

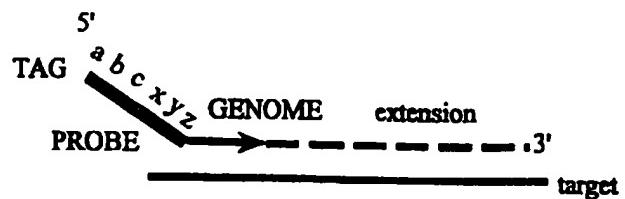
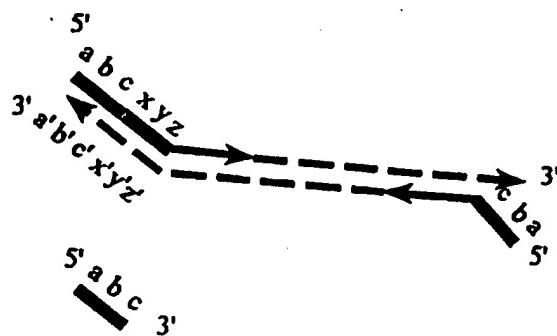
7/17

**FIGURE 7****SUBSTITUTE SHEET (RULE 26)**

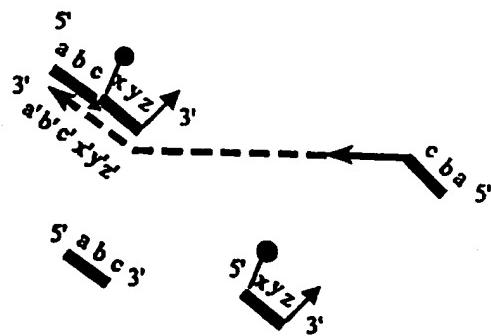
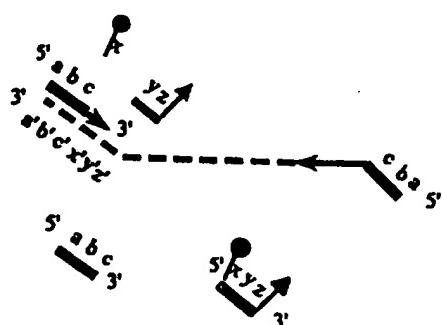
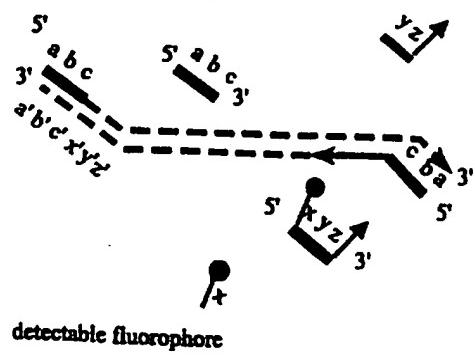
8/17

**FIGURE 8**

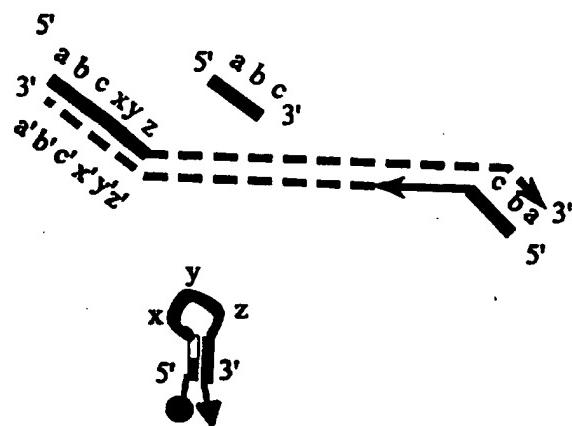
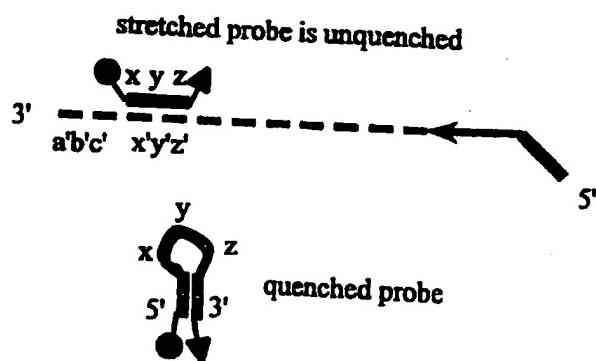
9/17

**FIGURE 9(a)****FIGURE 9(b)****SUBSTITUTE SHEET (RULE 26)**

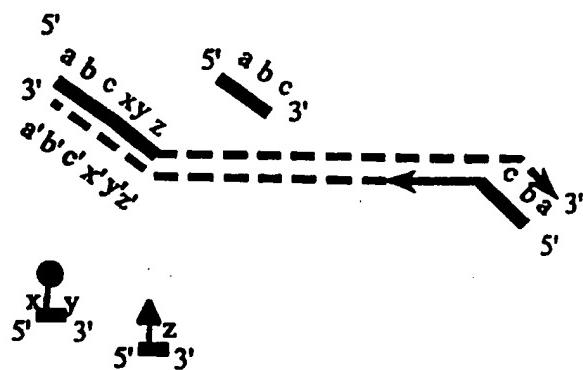
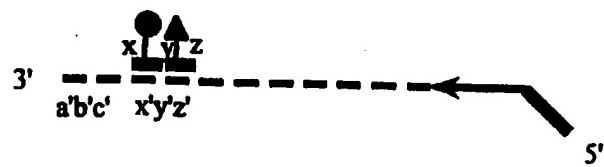
10/17

FIGURE 10(a)FIGURE 10(b)FIGURE 10(c)

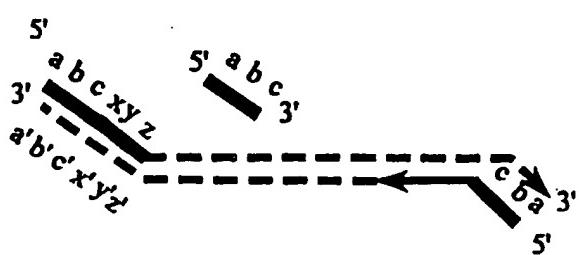
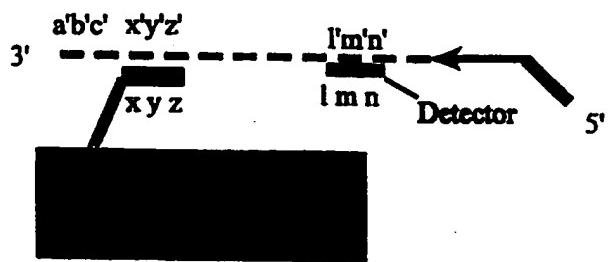
SUBSTITUTE SHEET (RULE 26)

**FIGURE 11(a)****FIGURE 11(b)**

12/17

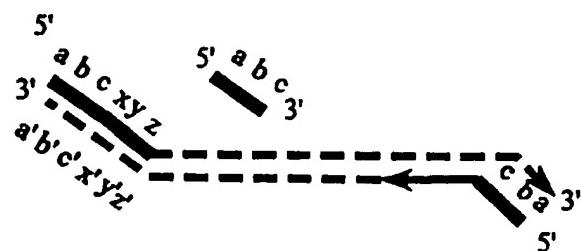
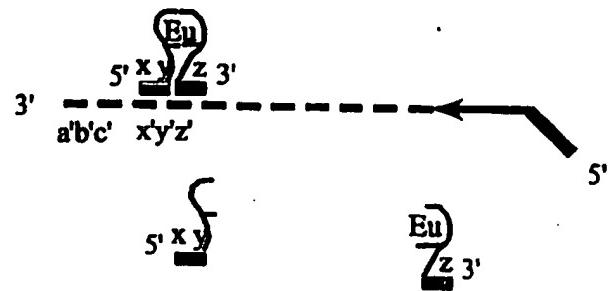
**FIGURE 12(a)****FIGURE 12(b)****SUBSTITUTE SHEET (RULE 26)**

13/17

FIGURE 13(a)FIGURE 13(b)

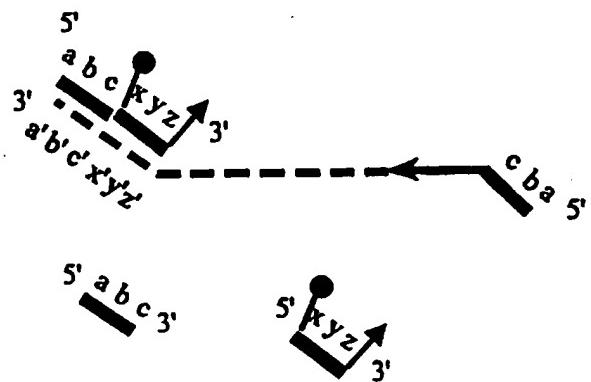
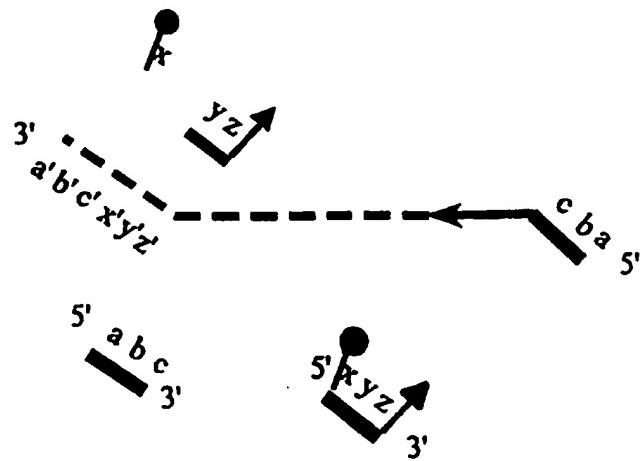
SUBSTITUTE SHEET (RULE 26)

14/17

FIGURE 14(a)FIGURE 14(b)

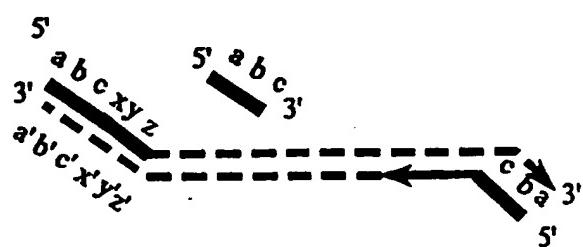
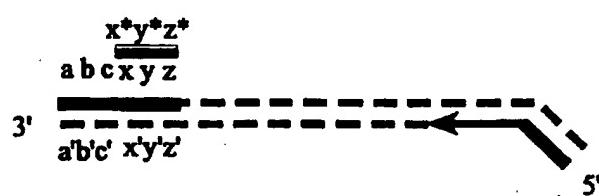
SUBSTITUTE SHEET (RULE 26)

15/17

FIGURE 15(a)FIGURE 15(b)

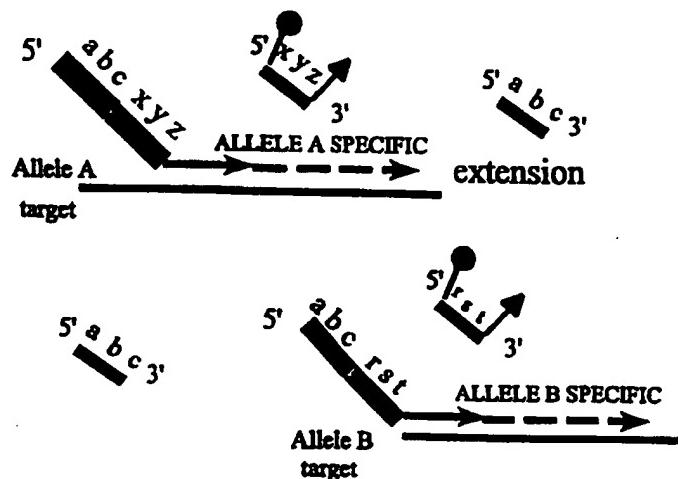
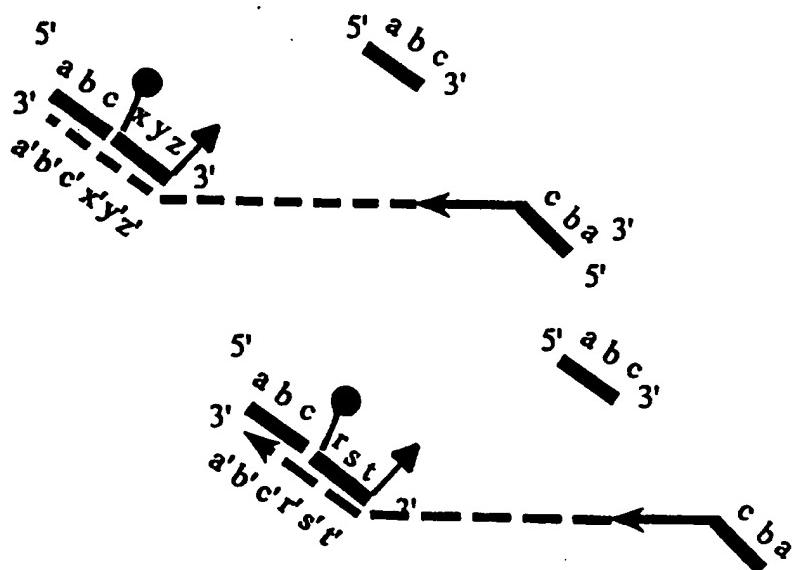
SUBSTITUTE SHEET (RULE 26)

16/17

FIGURE 16(a)FIGURE 16(b)

SUBSTITUTE SHEET (RULE 26)

17/17

FIGURE 17(a)FIGURE 17(b)

# INTERNATIONAL SEARCH REPORT

Int. Application No  
PCT/GB 97/01163

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 6 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 94 21820 A (HOPE CITY ;WALLACE ROBERT BRUCE (US)) 29 September 1994 see abstract ---	1,20,21
A	WO 95 08642 A (ZENECA LTD ;SAMMES PETER GEORGE (GB); GARMAN ANDREW JOHN (GB)) 30 March 1995 cited in the application see abstract ---	2,4,6,7
A	EP 0 678 581 A (BECTON DICKINSON CO) 25 October 1995 cited in the application see abstract ---	2,4,5

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

\* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*&\* document member of the same patent family

Date of the actual completion of the international search  26 August 1997	Date of mailing of the international search report  09.09.97
Name and mailing address of the ISA  European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	Authorized officer  Ceder, O

## INTERNATIONAL SEARCH REPORT

Int'l Application No  
PCT/GB 97/01163

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 639 647 A (TANABE SEIYAKU CO ; EIKEN CHEMICAL (JP)) 22 February 1995 cited in the application see abstract -----	1
A	EP 0 416 817 A (ICI PLC) 13 March 1991 see abstract -----	1

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

Int'l Application No
PCT/GB 97/01163

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9421820 A	29-09-94	AU 3933393 A	11-10-94
WO 9508642 A	30-03-95	AU 7662794 A EP 0720659 A GB 2283095 A JP 9503660 T	10-04-95 10-07-96 26-04-95 15-04-97
EP 0678581 A	25-10-95	US 5547861 A US 5593867 A AU 1502395 A BR 9501583 A CA 2145576 A CA 2145719 A EP 0678582 A JP 7289299 A JP 8038199 A SG 30350 A AU 1501995 A	20-08-96 14-01-97 26-10-95 14-11-95 19-10-95 19-10-95 25-10-95 07-11-95 13-02-96 01-06-96 18-04-96
EP 0639647 A	22-02-95	JP 7023800 A	27-01-95
EP 0416817 A	13-03-91	AT 144290 T AU 622349 B AU 6118490 A DE 69028891 D DE 69028891 T GB 2235689 A,B JP 3272686 A US 5525494 A	15-11-96 02-04-92 14-03-91 21-11-96 13-03-97 13-03-91 04-12-91 11-06-96